

Reconocimiento Óptico de Caracteres, aplicación móvil para el control de los medicamentos consumidos por los pacientes.

Gloria Diodati, Adrián Gómez, Marcela Martínez, Daniel Luna

^a Departamento de Informática en Salud, Hospital Italiano de Buenos Aires, Argentina

Resumen. Objetivos: El objetivo del trabajo es la descripción de las herramientas y algoritmos utilizados para la construcción de un prototipo mobile, con la capacidad de realizar un reconocimiento óptico de caracteres (OCR) de la imagen del empaque del medicamento obtenida mediante el uso de la cámara del dispositivo o su código de barras y obtener información farmacéutica.

Metodología: Implementación de un prototipo para la plataforma IOS utilizando el lenguaje de programación Objective C y C.

Resultados: Se implementó un prototipo capaz de reconocer texto en una imagen que incluye algoritmos de procesamiento de imágenes, algoritmos de procesamiento de texto y el estudio de técnicas algorítmicas para la manipulación y búsqueda de información en la base de datos.

Discusión: Las técnicas utilizadas debieron contemplar aquellas limitaciones propias de la plataforma: tiempo de procesamiento y capacidad de almacenamiento. Se decidió trabajar con aquellos algoritmos y estructuras que presentaron el mejor rendimiento en ambos aspectos, en relación con los resultados obtenidos.

Palabras Claves: Personal Health Record, Mobile Health, Optical Character Recognition, Image Processing.

1 Introducción

La automedicación es una actividad común en la sociedad, definimos automedicación como el uso de un producto farmacéutico para el tratamiento de una enfermedad o síntoma sin la prescripción de un profesional. La automedicación puede tener lugar mediante la utilización de medicamentos sobrantes, utilización de remedios caseros, compartir medicamentos con familiares o miembros del grupo social, o bien no cumplir las indicaciones médicas prolongando o disminuyendo el tratamiento [1-2].

La confección y verificación en conjunto con el paciente de una lista completa y correcta de la medicación que consume realmente, y la correspondiente actualización de su registro en la historia clínica electrónica, son elementos esenciales en la evaluación del paciente cuando es atendido en el hospital o en el ámbito ambulatorio y resultan imprescindibles para asegurar la coherencia con la nueva medicación que se le prescribe [3-6]. Bajo este concepto los portales personales de salud (PHR) son herramientas que colaboran en una correcta comunicación entre el médico y el paciente acerca de la información médica y posibilitan una participación activa en el autocuidado de la salud [7-9]. Estudios recientes muestran un crecimiento en el interés en este tipo de herramientas como elemento necesario para el empoderamiento de los pacientes, reportándose en la literatura estudios que demuestran un alto grado de satisfacción por parte de los pacientes usuarios de estas herramientas potenciando la comunicación con los profesionales de la salud [10-13].

El objetivo del trabajo es la descripción de las herramientas y algoritmos utilizados para la construcción de un prototipo mobile, con la capacidad de realizar un reconocimiento óptico de caracteres (OCR) de la imagen del empaque del medicamento obtenida mediante el uso de la cámara del dispositivo o su código de barras y obtener información farmacéutica que se pueda compartir con la historia clínica electrónica alimentando la lista de fármacos que el paciente consume.

2 Materiales y Métodos

El Hospital italiano de Buenos Aires tiene implementado un Sistema de Información en Salud desde el año 1998, e implementó en el año 2007 un Portal Personal de Salud con aproximadamente 118.000 usuarios enrolados. A mediados del año 2012 comenzó el desarrollo de una versión móvil del Portal Personal de Salud, que estuvo disponible para los usuarios en enero del año 2013 contando con 12290 usuarios [14-15].

La implementación del prototipo se realizó para plataforma IOS utilizando el lenguaje de programación Objective C y C, para el reconocimiento de texto en la imagen se utilizó la herramienta Tesseract-ios (Tesseract OCR) junto con las bibliotecas de Leptonica. Para la lectura del código de barras se utilizó ZBar iPhone SDK 1.2[16-18].

Se realizó una abstracción de la base de datos farmacológica de la institución obteniendo un modelo compacto con la información farmacéutica necesaria para la validación de los medicamentos y la consulta de información contextual.

La base de datos es de tipo relacional y fue creada con la herramienta SQLite[17]. El modelo de información consiste en datos referidos a monodrogas, productos comerciales y monografías.

Para cada registro de la Tabla Medicamento se consignaron los siguientes atributos:

- Nombre comercial del producto.

- Presentación.
- Nombre controlado de las monodrogas que lo componen: tal como figura en español en la Nomenclatura Internacional No Propietaria de la Organización Mundial de la Salud (INN)[18].

En la Tabla Monografía se consigna la siguiente información para cada monodroga:

- Farmacológica.
- Terapéutica.
- Nombre controlado de la monodroga.

Finalmente para poder realizar la búsqueda por código de barras, y relacionar las dos tablas mencionadas anteriormente entre si, los registros de la Tabla Producto presentan información sobre el nombre comercial, monodroga, presentación y aquellos datos concernientes al código de barras (Fig. 1).

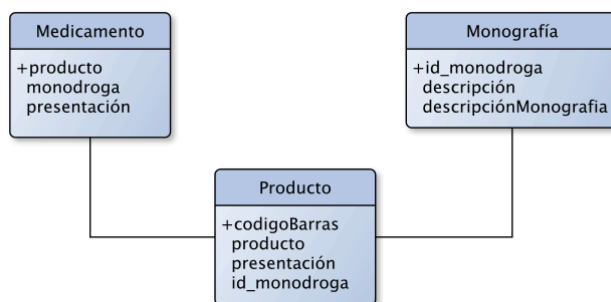


Figura 1:Modelo entidad relación

3 Resultados

En una primera etapa de procesamiento, se utilizó un algoritmo para ajustar el tamaño de la imagen a 640 x 640 píxeles. Una vez que la imagen fue reducida se procedió a estudiar la orientación en la que fue tomada y rotarla según correspondiera.

En una etapa posterior se convirtió la imagen a escala de grises y se aplicó un algoritmo de umbralización como método de segmentación. Este método consiste en convertir una imagen en escala de grises a una de dos niveles, con la finalidad de separar las letras del fondo del empaque. El proceso de umbralización se realizó con el valor medio de los píxeles de cada imagen como umbral.

Tanto el algoritmo de umbralización como la conversión a escala de grises y el cálculo del valor medio se realizó con la biblioteca OpenCV for iOS[19]. Dado que esta herramienta opera sobre estructuras matriciales y Objective C sobre imágenes, la conversión entre los tipos de datos se realizó mediante el framework Core Graphics.

Ya con la imagen preprocesada invocamos la herramienta Tesseract que va a ser la encargada de reconocer el texto. Esta herramienta fue entrenada previamente con los paquetes de datos correspondientes al lenguaje Español.

Con los datos obtenidos como resultado del proceso de reconocimiento procedemos a trabajar sobre el texto obtenido con el fin de identificar el nombre del producto, las principales monodrogas que lo componen y en que medida lo hacen. Como primera medida quitamos todos aquellos caracteres que no pertenecen al conjunto de símbolos alfanuméricos e ignoramos aquellas palabras de menos de tres caracteres.

Una vez que limpiamos el texto de potenciales caracteres erróneos continuamos con la identificación del producto y de la monodroga. Para lograrlo calculamos la Distancia de Levenshtein entre las palabras reconocidas y el contenido de los diccionarios de manera tal de poder seleccionar aquellas palabras cuyo costo de transformación sea el mínimo posible. Así logramos identificar tanto el producto como la monodroga.

En caso en que ninguna palabra reconocida tenga un costo de transformación aceptable, es decir, que a partir del texto reconocido no logremos identificar el medicamento, procedemos a invocar nuevamente a la herramienta para reconocer texto, pero esta vez, sobre la imagen obtenida como resultado de la primer etapa de preprocesamiento; es decir, aquella ajustada en tamaño y convertida a escala de grises.

Una vez identificado el producto y la monodroga que lo compone, buscamos el medicamento que cumple con dichas características en la base de datos (Fig. 2-3).

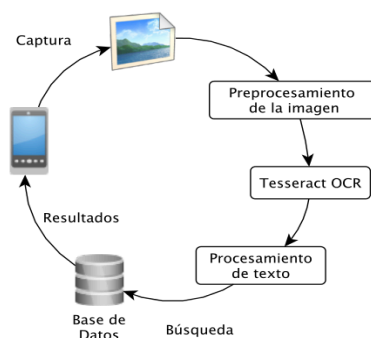


Figura 2: Diagrama de arquitectura para la primer alternativa.



Figura 3: Reconocimiento por empaque.

Otra alternativa que ofrece el programa es la de tomar una foto del código de barras del medicamento y, a partir del mismo, buscar la información deseada. Para reconocer dicho código utilizamos la herramienta ZBar. Una vez identificado el medicamento, buscamos en la base de datos las principales monodrogas y, a partir de ellas, la información del prospecto correspondiente (Fig. 4).



Figura 4: Reconocimiento por código de barras.

Finalmente, una tercera opción que ofrece el programa es la del ingreso del nombre comercial del medicamento de forma manual. Este método no necesita de los algoritmos de preprocesamiento de imagen pero sí de la validación del texto introducido. Al igual que en los casos anteriores, se realiza la consulta pertinente y se obtiene la información requerida.

Se realizó un proceso de test con varias imágenes de empaques de medicamentos, expuestas a distintas condiciones de luminosidad, en las que variaban tanto los colores como los materiales. Como resultado obtuvimos que el 74% de las imágenes fueron reconocidas en forma completa, el 14% fueron reconocidas con diferencias y el 12% no fueron reconocidas.

Entendemos como reconocimiento con diferencias aquellos casos en los que el medicamento fue reconocido pero se encontraron variaciones en algunas de las monodrogas que los componen. Dentro de las imágenes de los medicamentos que no fueron reconocidos podemos encontrar aquellas en las que el flash de la cámara o la luz artificial se vio reflejada en el empaque, produciendo un punto de mayor claridad y, consecuentemente, distorsionando las formas de las letras.

Dentro de las imágenes en las que se pudo reconocer el medicamento con algunas diferencias, encontramos aquellas en las que la tipografía del empaque no permitía distinguir correctamente las separaciones entre las letras en aquellas palabras que denotaban la particularidad del medicamento, por ejemplo: letra pequeña, cursiva, alineación no horizontal, etc.

Un experimento similar se llevó a cabo con imágenes de códigos de barras, en este caso se obtuvo una correcta detección del código numérico en todos los casos pero las deficiencias se encontraron en la completitud de la base de datos.

4 Discusión

La primer dificultad que se presentó fue la del reconocimiento del texto a partir de la imagen tomada con la cámara del dispositivo. Si bien contábamos con la herramienta Tesseract, las imágenes eran propensas a tener problemas de nitidez, luminosidad, falta de foco, etc. Estas perturbaciones en la imagen resultaron en un desempeño no deseado de la herramienta dado que originalmente fue diseñada para reconocer texto a partir de documentos escaneados y las técnicas de corrección incluidas en ella, no eran suficientes.

Para solucionar este inconveniente se procedió a estudiar distintos algoritmos de preprocesamiento de imágenes para remover las distintas distorsiones teniendo en consideración las limitaciones propias de un dispositivo móvil.

Entre los algoritmos y técnicas con los que se experimentó podemos nombrar: Bounding Box & Skew Angle Techniques[20], Hough Transform & Contour Algorithm[21] y Blob detection[22]. Estas técnicas fueron finalmente descartadas debido al alto costo tanto de memoria como de procesamiento del dispositivo, al mismo tiempo que los resultados no fueron significativamente buenos.

Finalmente se decidió trabajar con aquellos algoritmos que presentaron el mejor rendimiento en relación con los resultados: reducción del tamaño de la imagen, umbralización y conversión a escala de grises.

Con respecto a la reducción de la imagen, se estudió el impacto del tamaño en la precisión de los resultados obtenidos, suponiendo que a mayor tamaño mas acertados iban a ser. Esto no resultó como se esperaba ya que un mayor tamaño de imagen, implica una mayor cantidad de información a almacenar y procesar, lo que deriva en un aumento en la memoria utilizada y en el trabajo del procesador. Al mismo tiempo se notó que la calidad de los resultados, a partir de cierto tamaño, no mejoraba significativamente por lo que se decidió que 640 x 640 era una dimensión apropiada.

Durante el proceso de umbralización se priorizó que no se produjera una pérdida de información significativa, para lo cual, se utilizó como umbral el valor medio de los pixeles de cada imagen. De esta manera, cada una, se procesó con el valor mas apropiado posible.

La segunda dificultad a sortear fue la de quitar aquellos caracteres erróneos obtenidos del proceso de reconocimiento del texto de la imagen. A estos caracteres los podíamos encontrar:

- Aislados: con espacios en blanco a derecha y a izquierda.
- Agrupados con otros caracteres erróneos: formando una cadena invalida.
- En medio de una palabra: con letras a ambos lados.

Estas formas de encontrarlos introducían “ruido” a la hora de buscar en la base de datos por lo cual era necesario eliminarlos. Si bien se solucionó fácilmente ya que estas características permitían filtrarlos a través del conjunto de caracteres alfanuméricos, en algunos casos, se generaba una segmentación en la palabra produciendo una potencial pérdida de información. Para solucionar este inconveniente se calculó el "parecido" de la palabra segmentada con las palabras de los diccionarios mediante el cálculo de la Distancia de Levenshtein; de esta manera pudimos obtener aquella palabra cuyo "costo" de transformación era menor, es decir la mas parecida. Además, dependiendo de que diccionario proviniera dicha palabra, también pudimos saber si se trataba del nombre comercial de un producto o de una de las monodrogas que lo componía.

En caso en que ninguna palabra de las detectadas fuese parecida a algún nombre comercial o monodroga, se repite el procedimiento a partir de la etapa de reconocimiento con la herramienta Tesseract y con la imagen reducida y en escala de grises, como entrada. Se decidió reconocer el texto en la imagen umbralizada antes que en aquella en escala de grises ya que la mayor cantidad de resultados satisfactorios de la etapa de reconocimiento se obtuvieron a partir de la primer imagen por lo que el reconocimiento sobre la segunda no fue necesario.

Para lograr reconocer un medicamento mediante el código de barras se experimentó el funcionamiento de la herramienta ZBar. Dado que los resultados obtenidos a partir de esta herramienta fueron mas acertados que los de la herramienta Tesseract, no fue necesario preprocesar la imagen ni trabajar sobre el texto reconocido lo que resulta en un tiempo de ejecución menor al del método anterior.

Un inconveniente de esta alternativa es que solo aquellos medicamentos de venta libre llegan a las manos del consumidor con los datos del código de barras disponible.

Con respecto a la información del prospecto, resultó imperativo almacenar dicha información relacionada con los datos que pudiéramos obtener del empaque. Por un lado, necesitábamos de los nombres comerciales de todos aquellos medicamentos en venta, junto con sus distintas presentaciones y las monodrogas que los componen y, por el otro, la información de tipo farmacológica y terapéutica de cada monodroga.

En nuestro país existen varias empresas que mantienen publicaciones periódicas que contienen altas, bajas y modificaciones de los productos farmacéuticos comercializados en la región[23-24]. Estas publicaciones están destinadas exclusivamente a la gestión fármaco-contable de las farmacias de venta al público, las empresas mandatarias y las instituciones de salud. Dado que están diseñadas principalmente para la gestión contable, carecen de ciertos aspectos necesarios para nuestro uso; a saber:

- Identificación unívoca de productos: cada una de estas empresas mantiene un identificador propio para los productos farmacéuticos.
- El nombre comercial: Puede existir un producto que figure con nombres diferentes en cada base comercial. También es posible que un producto de reciente aparición en el mercado figure solamente en una de ellas y no en otra.
- Las presentaciones: La aparición en el mercado de nuevas presentaciones puede reflejarse en diferentes momentos en las distintas bases.
- La completitud del dominio: Aun si se sumaran los contenidos de las dos bases comerciales disponibles, muchos productos, entre ellos aquellos de utilización hospitalaria y gran cantidad de los productos importados, no estarían en ella.
- Las monodrogas que componen el producto.
- Las acciones farmacológicas.

Para solucionar estas deficiencias, se decidió crear una Tabla Medicamento, donde se logra identificar de forma inequívoca a cada producto y presentación y enlazar cada registro de la lista con la representación que tenga en una o mas bases de datos comerciales con el propósito de lograr la completitud de dominio necesaria. Los productos que no figuran en ninguna de las dos bases son incorporados de forma manual siguiendo las mismas rigurosas pautas de control establecidas para los que se obtienen las bases de datos comerciales.

Con respecto a las acciones farmacológicas de las monodrogas, las bases comerciales disponibles utilizan indistintamente términos que son ambiguos, vagos, duplicados e imprecisos. Por lo tanto, se decidió crear la Tabla Monografía con información farmacológica y terapéutica, relacionada con el término controlado de la monodroga. Entendemos como información farmacológica aquella relativa a la característica funcional de la monodroga, por ejemplo, “Bloqueante de los canales de calcio”; y terapéutica, aquella relativa al empleo en situaciones medicas específicas, por ejemplo “Antihipertensivo”.

5 Conclusión

En este trabajo se planteó como objetivo la obtención de información terapéutica y farmacológica de un medicamento mediante los datos obtenidos a partir de una imagen del empaque. Esto se logró a través del uso de herramientas de reconocimiento de caracteres, sin embargo, para obtener mejores resultados en esta etapa, fue necesario pre-procesar la imagen de entrada. Para seleccionar los algoritmos de pre-procesamiento mas adecuados a la plataforma de trabajo, se estudiaron distintas técnicas sin olvidar las propias limitaciones de la misma.

En cuanto a la información a mostrar, fue necesario crear una base de datos conteniendo los nombres comerciales, monodrogas y acciones terapéuticas y farmacológicas de los medicamentos. También fue necesario trabajar sobre el texto resultante del proceso de reconocimiento debido a que algunos caracteres obtenidos no pertenecían al conjunto de caracteres alfanuméricos, introduciendo "ruido" en el proceso de búsqueda.

Otra alternativa que se propuso fue la obtención de la información a partir de la imagen del código de barras del empaque. Esta opción fue mas simple de implementar dado que la herramienta utilizada para reconocer el código presenta resultados mas atinados que el primer método mencionado, por lo que no fue necesario pre-procesar la imagen ni trabajar sobre el texto reconocido. Al no realizar estas operaciones, el tiempo de procesamiento final disminuyó notablemente.

A futuro, se podrá incluir el reconocimiento de los medicamentos a través de la lectura de un código QR y se trabajara en la confeccion de un protocolo de investigacion para realizar un analisis de la eficiencia y eficacia del prototipo mediante el uso de metodologias formales de la investigacion.

Referencias

1. Tam vc, Knowles SR, . Frequency, type and clinical importance of medication history errores at admission. Systematic review. CMAJ 2005; 173:510.
2. Gleason KM, Groszek JM. Reconciliation of discrepancies in medication. Am J Health sys Pharm. 2004
3. Segall A. A community survey of self-medication activities. Med Care 1990; 28:301-10.
4. Wilkinson IF, Darby DN, Mant A. Self-care and self-medication: an evaluation of individuals' health care decisions. Med Care 1987; 25:965-78.
5. Blenkinsopp A, Bradley C. Patients, society and the increase in self-medication. BMJ 1996; 312:
6. Bradley C, Blenkinsopp A. Over-the-counter drugs: the future of self-medication. BMJ 1996; 312:835-7.
7. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. BMJ. 2007;335:330-3. [PMID: 17703042]
8. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J

9. Am Med Inform Assoc. 2006;13:121-6. [PMID: 16357345]
10. Ball MJ, Costin MY. The personal health record: consumers banking on their health. *Stud Health Technol Inform* 2008;134:35-46
11. Latham CE. Is there data to support the concept that educated, empowered patient have better outcomes? *J Am Soc Neprrol*.
12. Cimino JJ, Patel VL. The patient clinical information system. *Int J Med Inform* 2002.
- 13.
14. Otero C. MM, Luna D. Evaluación del perfil de uso del portal personal de salud del Hospital Italiano de Buenos Aires 2011.
15. Julieta Goldenberg MM, Enrique Stanziola, Adrián Gómez, Daniel Luna, Quirós. FGBd. Rediseño Centrado en el Usuarios de un Portal Personal de Salud 2012.
16. <http://code.google.com/p/tesseract-ocr>
17. <http://github.com/ldiquial/tesseract-ios-lib>
18. <http://zbar.sourceforge.net/index.html>
19. <http://www.sqlite.org/>
20. World Health Organization, International Nonproprietary Names (INN) for Pharmaceutical Substances, Cumulative List No. 9. 1996, Geneva, Switzerland. 885 pages.
21. <http://opencv.org/downloads.html>
22. <http://felix.abecassis.me/2011/10/opencv-bounding-box-skew-angle/>
23. Duan, Duc, Du, 2004, *Combining Hough Transform and Contour Algorithm for detecting Vehicles License-Plates*, Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong.
24. <http://areshopencv.blogspot.pt/2011/12/blob-detection-connected-component-pure.html>
25. <http://ar.kairosweb.com/> Accedido: julio 3, 2014
26. AlfaBetaEdiciones, *Manual Farmacéutico* <http://web.alfabeta.net/mfonline.xtp> Accedido: Julio 3, 2014

Dirección para correspondencia

Gloria Diodati: gloriadiodati@gmail.com

Adrián Gómez: adrian.gomez@hospitalitaliano.org.ar